Online Appendix
for
Market efficiency in the age of big data

Ian W. R. Martin*
London School of Economics and CEPR

Stefan Nagel†
University of Chicago, NBER, CEPR, and CESIfo

In this online appendix, we consider the case where an econometrician observes only a subset of the firm characteristics that are observable to investors. Specifically, we now assume that the econometrician's predictors are collected in the $N \times P$ matrix $\boldsymbol{C}$, where $P \leq J$ and $\text{rank}(\boldsymbol{C}) = P$.

Assumption OA.1. *$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{C} & \boldsymbol{M} \end{pmatrix}$ where $\boldsymbol{M}$ is an $N \times (J-P)$ matrix and $\boldsymbol{C}'\boldsymbol{M} = \boldsymbol{0}$. Furthermore, the variables in $\boldsymbol{C}$ have been scaled such that $\text{tr}(\boldsymbol{C}'\boldsymbol{C}) = NP$. (As $\text{tr}\,\boldsymbol{X}'\boldsymbol{X} = \text{tr}\,\boldsymbol{C}'\boldsymbol{C} + \text{tr}\,\boldsymbol{M}'\boldsymbol{M}$, it follows that $\text{tr}(\boldsymbol{M}'\boldsymbol{M}) = N(J-P)$.)*

We view the econometrician's predictive variables $\boldsymbol{C}$ as fixed prior to observing the data. The assumption that $\boldsymbol{C}'\boldsymbol{M} = 0$ is without loss of generality, as we can think of the investors using the predictor variables observed by the econometrician, $\boldsymbol{C}$, together with extra unobserved variables $\boldsymbol{M}$ that are residualized with respect to $\boldsymbol{C}$.

The econometrician regresses $\boldsymbol{r}_{t+1}$ on $\boldsymbol{C}$, obtaining a vector of cross-sectional regression coefficients

$$\boldsymbol{b}_{t+1} = (\boldsymbol{C}'\boldsymbol{C})^{-1}\,\boldsymbol{C}'\boldsymbol{r}_{t+1}\,. \tag{OA.1}$$

Under the rational expectations null,

$$\sqrt{N}\boldsymbol{b}_{t+1} \sim N\left(0, N(\boldsymbol{C}'\boldsymbol{C})^{-1}\right), \tag{OA.2}$$

and

$$\boldsymbol{b}_{t+1}'(\boldsymbol{C}'\boldsymbol{C})\boldsymbol{b}_{t+1} \sim \chi^2_P. \tag{OA.3}$$

*London School of Economics; i.w.martin@lse.ac.uk.
†University of Chicago, Booth School of Business; stefan.nagel@chicagobooth.edu.

As we want to characterize the properties of the econometrician's test under asymptotics where $N, P \to \infty$ and $P/N \to \phi > 0$, it is more convenient if we let the econometrician consider a scaled version of this test statistic:

$$T_{re} \equiv \frac{\boldsymbol{b}_{t+1}' \boldsymbol{C}' \boldsymbol{C} \boldsymbol{b}_{t+1} - P}{\sqrt{2P}}. \tag{OA.4}$$

Under the econometrician's rational expectations null, we would have, asymptotically,

$$T_{re} \xrightarrow{d} N(0,1) \quad \text{as } N, P \to \infty,\ P/N \to \phi > 0. \tag{OA.5}$$

But the actual asymptotic distribution of T_{re} is influenced by the components of returns involving $\bar{\boldsymbol{e}}_t$ and $\boldsymbol{g}$ in (8) in the main text.

These alter the asymptotic distribution and may lead the rejection probabilities of a test using $N(0,1)$ critical values based on (OA.5), or χ^2 critical values based on (OA.3), to differ from the nominal size of the test. As in our main analysis, we assume that $\boldsymbol{g}$ is drawn from the prior distribution. Here, we define

$$\boldsymbol{\Sigma}_{re} = (\boldsymbol{C}'\boldsymbol{C})^{-1} \quad \text{and} \quad \boldsymbol{\Sigma}_b = \mathbb{E}\left(\boldsymbol{b}_{t+1}\boldsymbol{b}_{t+1}'\right),$$

and

$$\zeta_{i,t} = 1 + \frac{1}{t + \frac{J}{N\theta\lambda_i}} \tag{OA.6}$$

are the eigenvalues of $\boldsymbol{\Sigma}_b \boldsymbol{\Sigma}_{re}^{-1}$ that control the asymptotic behavior of T_{re} with limiting mean and variance of

$$\mu = \lim_{P\to\infty} \frac{1}{P} \sum_{i=1}^{P} \zeta_{i,t} \quad \text{and} \quad \sigma^2 = \lim_{P\to\infty} \frac{1}{P} \sum_{i=1}^{P} \zeta_{i,t}^2 - \mu^2 .$$

By the "Big Data" Assumption 5 in the main text, we have $1 < \mu < 2$ and $1 < \sqrt{\mu^2 + \sigma^2} < 2$ for all $t \geq 1$.

Proposition OA.1. *If returns are generated according to (8) in the main text, then in the large N, P limit*

$$\frac{\boldsymbol{b}_{t+1}' \boldsymbol{C}' \boldsymbol{C} \boldsymbol{b}_{t+1} - \sum_{i=1}^{P} \zeta_{i,t}}{\sqrt{2 \sum_{i=1}^{P} \zeta_{i,t}^2}} \xrightarrow{d} N(0,1) .$$

It follows that the test statistic T_{re} satisfies

$$\frac{T_{re}}{\sqrt{\mu^2 + \sigma^2}} - \frac{\mu - 1}{\sqrt{2(\mu^2 + \sigma^2)}} \sqrt{P} \xrightarrow{d} N(0,1)$$

where $1 < \mu < 2$ *and* $1 < \sqrt{\mu^2 + \sigma^2} < 2$.

We can therefore think of T_{re} as a multiple of a standard Normal random variable plus a term of order $\sqrt{P}$:

$$T_{re} \approx \sqrt{\mu^2 + \sigma^2}\, N(0,1) + \frac{\mu - 1}{\sqrt{2}} \sqrt{P}\,. \tag{OA.7}$$

This result is similar to the corresponding result in our main analysis, but with $\sqrt{P}$ taking the place of $\sqrt{J}$.

Proposition OA.2. *In a test of return predictability based on the rational expectations null* (OA.5)*, we would have, for any critical value* c_α *and at any time* t*,*

$$\mathbb{P}\left(T_{re} > c_\alpha\right) \to 1 \text{ as } N, P \to \infty, P/N \to \phi\,.$$

More precisely, for any fixed $t > 0$*, the probability that the test fails to reject declines exponentially fast as* N *and* P *increase, at a rate that is determined by* μ*,* σ*, and* ϕ*:*

$$\lim_{N\to\infty} -\frac{1}{N} \log \mathbb{P}\left(T_{re} < c_\alpha\right) = \frac{(\mu - 1)^2 \phi}{4\left(\mu^2 + \sigma^2\right)}\,, \tag{OA.8}$$

for any critical value c_α*.*

Thus, as in our main analysis, in-sample predictability tests lose their economic meaning when P is not small relative to N, in the sense that rejection of the no-predictability can be likely even without risk premia or mispricing.

For out-of-sample tests, we obtain the following result.

Proposition OA.3. *If returns are generated according to (8) in the main text and* $r_{OOS,t+1} = \frac{1}{N}\boldsymbol{r}'_{t+1}\boldsymbol{C}\boldsymbol{b}_{s+1}$ *with* $s \neq t$*, then*

$$\mathbb{E}\, r_{OOS,t+1} = 0$$

and, in the large N, J, P *limit,*

$$\frac{r_{OOS,t+1}}{\sqrt{\sum_{j=1}^{P} \zeta_{j,s}\zeta_{i,t}}} \xrightarrow{d} N(0,1)\,.$$

Now suppose the characteristics are associated with a predictable component $\boldsymbol{X}\boldsymbol{\gamma}_x$ for some vector $\boldsymbol{\gamma}_x$ that represents risk premia. In this case, adding this component to the returns in (8) in the main text, we get

$$\boldsymbol{r}_{t+1} = \boldsymbol{X}\boldsymbol{\gamma}_x + \boldsymbol{X}(\boldsymbol{I} - \boldsymbol{\Gamma}_t)\boldsymbol{g} - \boldsymbol{X}\boldsymbol{\Gamma}_t(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\bar{\boldsymbol{e}}_t + \boldsymbol{e}_{t+1}, \tag{OA.9}$$

The econometrician wants to estimate to what extent the characteristics she observes (i.e., $\boldsymbol{C}$) are associated with risk premia. Projecting $\boldsymbol{X}$ on $\boldsymbol{C}$ we can decompose $\boldsymbol{X} = \boldsymbol{C}\boldsymbol{B} + \boldsymbol{E}$, where $\boldsymbol{E}'\boldsymbol{C} = \boldsymbol{0}$ and rewrite returns as

$$\boldsymbol{r}_{t+1} = \boldsymbol{C}\boldsymbol{\gamma} + \boldsymbol{E}\boldsymbol{\gamma}_x + \boldsymbol{X}(\boldsymbol{I} - \boldsymbol{\Gamma}_t)\boldsymbol{g} - \boldsymbol{X}\boldsymbol{\Gamma}_t(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\bar{\boldsymbol{e}}_t + \boldsymbol{e}_{t+1} \tag{OA.10}$$

where we have defined $\boldsymbol{\gamma} = \boldsymbol{B}\boldsymbol{\gamma}_x$.

Given the returns (OA.10). and using results from Proposition OA.3, it is straightforward to show that

$$\frac{1}{N}\boldsymbol{\gamma}'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{\gamma} = \mathbb{E}\, r_{OOS,t+1}. \tag{OA.11}$$

A Proofs

We first recall some notation and some basic facts that we will exploit throughout this appendix. We will use the eigendecomposition

$$\frac{1}{N}\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}' \tag{OA.12}$$

and the definition

$$\boldsymbol{\Gamma}_t = \boldsymbol{Q}\left(\boldsymbol{I} + \frac{J}{N\theta t}\boldsymbol{\Lambda}^{-1}\right)^{-1}\boldsymbol{Q}'. \tag{OA.13}$$

It will be convenient for future use to note that (OA.12) and (OA.13) imply that

$$\boldsymbol{I} - \boldsymbol{\Gamma}_t = \boldsymbol{Q}\left(\boldsymbol{I} + \frac{N\theta t}{J}\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{Q}' \tag{OA.14}$$

and

$$\frac{\theta t}{J}\left(\boldsymbol{I} - \boldsymbol{\Gamma}_t\right)\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{\Gamma}_t\,. \tag{OA.15}$$

We will repeatedly exploit the fact that $\boldsymbol{Q}$ is orthogonal (that is, $\boldsymbol{Q}\boldsymbol{Q}' = \boldsymbol{Q}'\boldsymbol{Q} = \boldsymbol{I}$) and $\boldsymbol{\Lambda}$ is diagonal. It follows that $\boldsymbol{\Gamma}_t$ is also diagonal. Note further that diagonal matrices commute.

Assumption OA.1 implies that

$$\frac{1}{N}\boldsymbol{X}'\boldsymbol{X} = \begin{pmatrix} \frac{1}{N}\boldsymbol{C}'\boldsymbol{C} & 0 \\ 0 & \frac{1}{N}\boldsymbol{M}'\boldsymbol{M} \end{pmatrix}$$

is block-diagonal (where zeros indicate conformable matrices of zeros). Eigendecomposing $\frac{1}{N}\boldsymbol{C}'\boldsymbol{C} = \boldsymbol{Q}_C\boldsymbol{\Lambda}_C\boldsymbol{Q}_C'$ and $\frac{1}{N}\boldsymbol{M}'\boldsymbol{M} = \boldsymbol{Q}_M\boldsymbol{\Lambda}_M\boldsymbol{Q}_M'$ where $\boldsymbol{Q}_i'\boldsymbol{Q}_i = \boldsymbol{I}$ for $i = C, M$, we have

$$\frac{1}{N}\boldsymbol{X}'\boldsymbol{X} = \underbrace{\begin{pmatrix}\boldsymbol{Q}_C & 0\\ 0 & \boldsymbol{Q}_M\end{pmatrix}}_{\boldsymbol{Q}}\underbrace{\begin{pmatrix}\boldsymbol{\Lambda}_C & 0\\ 0 & \boldsymbol{\Lambda}_M\end{pmatrix}}_{\boldsymbol{\Lambda}}\underbrace{\begin{pmatrix}\boldsymbol{Q}'_C & 0\\ 0 & \boldsymbol{Q}'_M\end{pmatrix}}_{\boldsymbol{Q}'}.$$

It follows that $\boldsymbol{\Gamma}_t = \begin{pmatrix}\boldsymbol{\Gamma}_{C,t} & 0\\ 0 & \boldsymbol{\Gamma}_{M,t}\end{pmatrix}$ is block-diagonal, with $\boldsymbol{\Gamma}_{i,t} = \boldsymbol{Q}_i\left(\boldsymbol{I} + \frac{J}{N\theta t}\boldsymbol{\Lambda}_i^{-1}\right)^{-1}\boldsymbol{Q}'_i$ and $\boldsymbol{I} - \boldsymbol{\Gamma}_{i,t} = \boldsymbol{Q}_i\left(\boldsymbol{I} + \frac{N\theta t}{J}\boldsymbol{\Lambda}_i\right)^{-1}\boldsymbol{Q}'_i$, where $\boldsymbol{Q}_i$ is orthogonal and $\boldsymbol{\Lambda}_i$ is diagonal for $i = C, M$.

Lastly, Assumption 4 in the main text implies that (i) $\boldsymbol{\Sigma}_g = (\theta/J)\boldsymbol{I}$, (ii) $\mathbb{E}\,\bar{\boldsymbol{e}}_t\boldsymbol{e}'_t = \mathbb{E}\,\boldsymbol{e}_t\bar{\boldsymbol{e}}'_t = \mathbb{E}\,\bar{\boldsymbol{e}}_t\bar{\boldsymbol{e}}'_t = \frac{1}{t}\boldsymbol{I}$, (iii) $\mathbb{E}\,\boldsymbol{e}_t\boldsymbol{e}'_t = \boldsymbol{I}$; and, for $s < t$, (iv) $\mathbb{E}\,\bar{\boldsymbol{e}}_t\boldsymbol{e}'_{s+1} = (1/t)\boldsymbol{I}$, (v) $\mathbb{E}\,\boldsymbol{e}_{t+1}\bar{\boldsymbol{e}}'_s = 0$, and (vi), $\mathbb{E}\,\bar{\boldsymbol{e}}_t\bar{\boldsymbol{e}}'_s = (1/t)\boldsymbol{I}$.

Proof of Proposition OA.1. We form $\boldsymbol{b}_{t+1} = (\boldsymbol{C}'\boldsymbol{C})^{-1}\boldsymbol{C}'\boldsymbol{r}_{t+1}$ and then look at the in-sample return $\boldsymbol{r}'_{t+1}\boldsymbol{C}(\boldsymbol{C}'\boldsymbol{C})^{-1}\boldsymbol{C}'\boldsymbol{r}_{t+1} = \boldsymbol{b}'_{t+1}\boldsymbol{C}'\boldsymbol{C}\boldsymbol{b}_{t+1}$. To think about the distribution of this quantity, we first need to understand $\boldsymbol{b}_{t+1}$ itself. It is a zero mean Normal random vector, and as $\mathbb{E}\,\boldsymbol{r}_{t+1}\boldsymbol{r}'_{t+1} = \frac{\theta}{J}\boldsymbol{X}(\boldsymbol{I} - \boldsymbol{\Gamma}_t)\boldsymbol{X}' + \boldsymbol{I}$ by Proposition 2 in the main text we have

$$\begin{aligned}
\mathbb{E}\,\boldsymbol{b}_{t+1}\boldsymbol{b}'_{t+1} &= (\boldsymbol{C}'\boldsymbol{C})^{-1}\boldsymbol{C}'\,\mathbb{E}\,\boldsymbol{r}_{t+1}\boldsymbol{r}'_{t+1}\boldsymbol{C}(\boldsymbol{C}'\boldsymbol{C})^{-1}\\
&= \frac{\theta}{J}(\boldsymbol{C}'\boldsymbol{C})^{-1}\boldsymbol{C}'\boldsymbol{X}(\boldsymbol{I} - \boldsymbol{\Gamma}_t)\boldsymbol{X}'\boldsymbol{C}(\boldsymbol{C}'\boldsymbol{C})^{-1} + (\boldsymbol{C}'\boldsymbol{C})^{-1}\\
&= \frac{\theta}{J}(\boldsymbol{I} - \boldsymbol{\Gamma}_{C,t}) + (\boldsymbol{C}'\boldsymbol{C})^{-1},
\end{aligned}$$

where we use the fact that $\boldsymbol{C}'\boldsymbol{X} = \begin{pmatrix}\boldsymbol{C}'\boldsymbol{C} & \boldsymbol{0}\end{pmatrix}$ in the last line. In terms of the eigendecomposition,

$$N\,\mathbb{E}\,\boldsymbol{b}_{t+1}\boldsymbol{b}'_{t+1} = \boldsymbol{Q}_C\underbrace{\left[\left(\frac{J}{N\theta}\boldsymbol{I} + t\boldsymbol{\Lambda}_C\right)^{-1} + \boldsymbol{\Lambda}_C^{-1}\right]}_{\boldsymbol{\Omega}_{C,t}}\boldsymbol{Q}'_C.$$

If we define $\boldsymbol{u}_{t+1} = \sqrt{N}\boldsymbol{\Omega}_C^{-1/2}\boldsymbol{Q}'_C\boldsymbol{b}_{t+1}$, then $\sqrt{N}\boldsymbol{b}_{t+1} = \boldsymbol{Q}_C\boldsymbol{\Omega}_C^{1/2}\boldsymbol{u}_{t+1}$ and $\boldsymbol{u}_{t+1}$ is standard Normal: $\mathbb{E}\,\boldsymbol{u}_{t+1}\boldsymbol{u}'_{t+1} = N\boldsymbol{\Omega}_C^{-1/2}\boldsymbol{Q}'_C\frac{1}{N}\boldsymbol{Q}_C\boldsymbol{\Omega}_C\boldsymbol{Q}'_C\boldsymbol{Q}_C\boldsymbol{\Omega}_C^{-1/2} = \boldsymbol{I}$. We can then write

$$\begin{aligned}
\boldsymbol{b}'_{t+1}\boldsymbol{C}'\boldsymbol{C}\boldsymbol{b}_{t+1} &= \underbrace{\boldsymbol{u}'_{t+1}\boldsymbol{\Omega}_C^{1/2}\boldsymbol{Q}'_C}_{\sqrt{N}\boldsymbol{b}'_{t+1}}\underbrace{\boldsymbol{Q}_C\boldsymbol{\Lambda}_C\boldsymbol{Q}'_C}_{\frac{1}{N}\boldsymbol{C}'\boldsymbol{C}}\underbrace{\boldsymbol{Q}_C\boldsymbol{\Omega}_C^{1/2}\boldsymbol{u}_{t+1}}_{\sqrt{N}\boldsymbol{b}_{t+1}}\\
&= \boldsymbol{u}'_{t+1}\boldsymbol{\Omega}_C^{1/2}\boldsymbol{\Lambda}_C\boldsymbol{\Omega}_C^{1/2}\boldsymbol{u}_{t+1}\\
&= \boldsymbol{u}'_{t+1}\boldsymbol{\Omega}_C\boldsymbol{\Lambda}_C\boldsymbol{u}_{t+1}.
\end{aligned}$$

The last line exploits the fact that $\boldsymbol{\Omega}_C^{1/2}$ and $\boldsymbol{\Lambda}_C$ commute, as they are diagonal.

It follows that

$$\boldsymbol{b}_{t+1}'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{b}_{t+1} = \sum_{i=1}^{P} \zeta_{i,t} u_i^2, \tag{OA.16}$$

where $\zeta_{i,t}$ are the diagonal entries of the diagonal matrix $\boldsymbol{\Omega}_C\boldsymbol{\Lambda}_C$ and u_i are independent $N(0,1)$ random variables (the entries of $\boldsymbol{u}_{t+1}$). Explicitly,

$$\zeta_{i,t} = \omega_{i,t}\lambda_i = \frac{\lambda_i}{t\lambda_i + \frac{J}{\theta N}} + 1. \tag{OA.17}$$

As $\lambda_i > 0$ by positive definiteness of $\boldsymbol{X}'\boldsymbol{X}$, it follows that for $t \geq 1$, $\zeta_{i,t} \in (1,2)$. Moreover, as $\lim_{J,N\to\infty} \frac{J}{N} = \psi > 0$ and (by Assumption (5) in the main text) $\lambda_i > \varepsilon$, $\zeta_{i,t}$ is uniformly bounded away from 1 and 2. It follows that $\mu \in (1,2)$ and $\sqrt{\mu^2+\sigma^2} \in (1,2)$.

We will apply Lyapunov's version of the central limit theorem to $\sum_{i=1}^{P} \zeta_{i,t} u_i^2$, which here requires that for some $\delta > 0$

$$\lim_{N,J\to\infty} \frac{1}{s_P^{2+\delta}} \sum_{i=1}^{P} \zeta_{i,t}^{2+\delta}\, \mathbb{E}\left[\left|u_i^2 - 1\right|^{2+\delta}\right] = 0 \quad \text{where} \quad s_P^2 = 2\sum_{i=1}^{P} \zeta_{i,t}^2.$$

It is enough to show that this holds for $\delta = 2$. But as $\mathbb{E}\left[(u_i^2-1)^4\right] = 60$ and $\zeta_{i,t} \in (1,2)$,

$$\frac{1}{s_P^4} \sum_{i=1}^{P} \zeta_{i,t}^4\, \mathbb{E}\left[\left|u_i^2 - 1\right|^4\right] = \frac{60\sum_{i=1}^{P} \zeta_{i,t}^4}{\left(2\sum_{i=1}^{P} \zeta_{i,t}^2\right)^2} \leq \frac{960P}{4P^2} \to 0 \quad \text{as } P \to \infty\,,$$

as required. Therefore the central limit theorem applies for $\boldsymbol{b}_{t+1}'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{b}_{t+1} = \sum_{i=1}^{P} \zeta_{i,t} u_i^2$ after appropriate standardization by mean and variance, which (as the u_i are IID standard Normal) are $\sum_{i=1}^{P} \zeta_{i,t}$ and $2\sum_{i=1}^{P} \zeta_{i,t}^2$, respectively. Thus we have

$$T_b \equiv \frac{\boldsymbol{b}_{t+1}'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{b}_{t+1} - \sum_{i=1}^{P} \zeta_{i,t}}{\sqrt{2\sum_{i=1}^{P} \zeta_{i,t}^2}} \xrightarrow{d} N(0,1)\,.$$

The remaining results follow immediately. □

Proof of Proposition OA.2. The first statement follows from the second. To prove the second, note that Proposition OA.1 implies that

$$\mathbb{P}\left(T_{re}<c_\alpha\right)=\mathbb{P}\left(\frac{T_{re}}{\sqrt{\mu^2+\sigma^2}}-\frac{\mu-1}{\sqrt{2\left(\mu^2+\sigma^2\right)}}\sqrt{P}<\frac{c_\alpha}{\sqrt{\mu^2+\sigma^2}}-\frac{\mu-1}{\sqrt{2\left(\mu^2+\sigma^2\right)}}\sqrt{P}\right)$$
$$\to\Phi\left(\frac{c_\alpha}{\sqrt{\mu^2+\sigma^2}}-\frac{\mu-1}{\sqrt{2\left(\mu^2+\sigma^2\right)}}\sqrt{P}\right),$$

where $\Phi(\cdot)$ denotes the standard Normal cumulative distribution function. The result follows from the well-known inequalities $\frac{e^{-x^2/2}}{\left|x+\frac{1}{x}\right|\sqrt{2\pi}}<\Phi(x)<\frac{e^{-x^2/2}}{|x|\sqrt{2\pi}}$, which hold for $x<0$. □

Proof of Proposition OA.3. We first show that $\mathbb{E}\, r_{OOS,t+1}=0$. As $\boldsymbol{b}_{s+1}=\left(\boldsymbol{C}'\boldsymbol{C}\right)^{-1}\boldsymbol{C}'\boldsymbol{r}_{s+1}$, we have

$$\mathbb{E}\left[\boldsymbol{r}_{t+1}\left(\boldsymbol{C}\boldsymbol{b}_{s+1}\right)'\right]=\mathbb{E}\left[\boldsymbol{r}_{t+1}\boldsymbol{r}_{s+1}'\right]\boldsymbol{C}\left(\boldsymbol{C}'\boldsymbol{C}\right)^{-1}\boldsymbol{C}'.$$

We will show that $\mathbb{E}\left[\boldsymbol{r}_{t+1}\boldsymbol{r}_{s+1}'\right]=0$ when $s\neq t$; in other words, all non-contemporaneous autocorrelations and cross-correlations are zero. Henceforth we assume that $s<t$ without loss of generality. From equation (8) in the main text,

$$\mathbb{E}\left[\boldsymbol{r}_{t+1}\boldsymbol{r}_{s+1}'\right]=\frac{\theta}{J}\boldsymbol{X}(\boldsymbol{I}-\boldsymbol{\Gamma}_t)(\boldsymbol{I}-\boldsymbol{\Gamma}_s)\boldsymbol{X}'+\frac{1}{t}\boldsymbol{X}\boldsymbol{\Gamma}_t(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{\Gamma}_s\boldsymbol{X}'-\frac{1}{t}\boldsymbol{X}\boldsymbol{\Gamma}_t(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'.$$

This expression can be rearranged as

$$\mathbb{E}\left[\boldsymbol{r}_{t+1}\boldsymbol{r}_{s+1}'\right]=\frac{1}{t}\boldsymbol{X}\left[\frac{\theta t}{J}(\boldsymbol{I}-\boldsymbol{\Gamma}_t)\boldsymbol{X}'\boldsymbol{X}-\boldsymbol{\Gamma}_t\right](\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{I}-\boldsymbol{\Gamma}_s)\boldsymbol{X}'.$$

As $\frac{\theta t}{J}(\boldsymbol{I}-\boldsymbol{\Gamma}_t)\boldsymbol{X}'\boldsymbol{X}=\boldsymbol{\Gamma}_t$ by equation (30) in the main text, the term in square brackets on the right-hand side vanishes, and the result follows.

We now turn to the asymptotic distribution. As $\boldsymbol{r}_{t+1}'\boldsymbol{C}\boldsymbol{b}_{s+1}=\boldsymbol{b}_{t+1}'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{b}_{s+1}$, we want to understand the behavior of $\boldsymbol{b}_{t+1}\boldsymbol{C}'\boldsymbol{C}\boldsymbol{b}_{s+1}$ where $s<t$. Defining $\boldsymbol{v}_{t+1}=\sqrt{N}\boldsymbol{\Omega}_{C,t}^{-1/2}\boldsymbol{Q}_C'\boldsymbol{b}_{t+1}$, as in the proof of Proposition OA.1, we have $\sqrt{N}\boldsymbol{b}_{t+1}=\boldsymbol{Q}_C\boldsymbol{\Omega}_{C,t}^{1/2}\boldsymbol{v}_{t+1}$ and $\mathbb{E}\,\boldsymbol{v}_{t+1}\boldsymbol{v}_{t+1}'=\boldsymbol{I}$. We also have $\mathbb{E}\,\boldsymbol{v}_{s+1}\boldsymbol{v}_{t+1}'=0$ whenever $s\neq t$ (because, as shown above, $\mathbb{E}\,\boldsymbol{r}_{s+1}\boldsymbol{r}_{t+1}'=0$ and hence $\mathbb{E}\,\boldsymbol{b}_{s+1}\boldsymbol{b}_{t+1}'=0$). Thus $\boldsymbol{v}_{t+1}$ and $\boldsymbol{v}_{s+1}$ are independent standard Normal random vectors. We have

$$\boldsymbol{b}_{t+1}'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{b}_{s+1}=\boldsymbol{v}_{t+1}'\boldsymbol{\Omega}_{C,t}^{1/2}\boldsymbol{Q}_C'\boldsymbol{Q}_C\boldsymbol{\Lambda}_C\boldsymbol{Q}_C'\boldsymbol{Q}_C\boldsymbol{\Omega}_{C,s}^{1/2}\boldsymbol{v}_{s+1}=\boldsymbol{v}_{t+1}'\boldsymbol{\Omega}_{C,t}^{1/2}\boldsymbol{\Lambda}_C\boldsymbol{\Omega}_{C,s}^{1/2}\boldsymbol{v}_{s+1}.$$

As $\boldsymbol{\Omega}_{C,t}^{1/2}\boldsymbol{\Lambda}_C\boldsymbol{\Omega}_{C,s}^{1/2}$ is a $P\times P$ diagonal matrix with ith diagonal entry $\sqrt{\zeta_{i,t}\zeta_{i,s}}$, we can write

$$\boldsymbol{b}_{t+1}'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{b}_{s+1}=\sum_{i=1}^{P}\sqrt{\zeta_{i,t}\zeta_{i,s}}\,w_i,$$

where w_i denotes the product of the ith entries of $\boldsymbol{v}_{t+1}$ and $\boldsymbol{v}_{s+1}$. The w_i are independent of each other, and each is the product of two independent standard Normal random variables. Therefore each w_i has zero mean and unit variance.

We wish to apply Lyapunov's version of the central limit theorem to $\sum_{i=1}^{P} \sqrt{\zeta_{i,t}\zeta_{i,s}}w_i$, which here requires that for some $\delta > 0$

$$\lim_{N,J,P\to\infty} \frac{1}{s_P^{2+\delta}} \sum_{p=1}^{P} \mathbb{E}\left[\left|\sqrt{\zeta_{i,t}\zeta_{i,s}}w_i\right|^{2+\delta}\right] = 0 \quad \text{where} \quad s_P^2 = \sum_{p=1}^{P} \zeta_{i,t}\zeta_{i,s}.$$

It is enough to show that this holds when $\delta = 2$. In this case, as the fourth moment of a standard Normal random variable equals 3, we have $\mathbb{E}\left[w_i^4\right] = 9$, and so indeed

$$\frac{1}{s_P^4} \sum_{p=1}^{P} \mathbb{E}\left[\left|\sqrt{\zeta_{i,t}\zeta_{i,s}}w_i\right|^{4}\right] = \frac{9\sum_{p=1}^{P} \zeta_{i,t}^2\zeta_{i,s}^2}{\left(\sum_{p=1}^{P} \zeta_{i,s}\zeta_{i,t}\right)^2} \leq \frac{144P}{P^2} \to 0 \quad \text{as } P \to \infty.$$

(The inequality follows because $\zeta_{i,t} \in (1,2)$ for all i and $t \geq 1$.) Hence the central limit theorem applies for $\sum_{i=1}^{P} \sqrt{\zeta_{i,t}\zeta_{i,s}}w_i$ after appropriate standardization by mean and variance, which are 0 and $\sum_{i=1}^{P} \zeta_{i,t}\zeta_{i,s}$, respectively. Thus

$$\frac{\boldsymbol{b}_{t+1}'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{b}_{s+1}}{\sqrt{\sum_{i=1}^{P} \zeta_{i,t}\zeta_{i,s}}} \xrightarrow{d} N(0,1)\,. \qquad \square$$